Wednesday, November 1 **Fall 2023**

Section 7: Poisson Processes, Joint Distributions, and Covariance

Srihari Ganesh

Based on section note formatting template by Rachel Li and Ginnie Ma '23

Forms

- Attendance form: http://bit.ly/110attend
- Feedback form: https://bit.ly/SrihariFeedback

Attendance form

Section feedback form

1 Summary

1.1 Moment generating functions

Definition 1 (Moments)**.** For a random variable X, the $\mathbf{n^{th}}$ **moment** is $E(X^n)$.

Definition 2 (Moment Generating Function)**.** For a random variable X, the **moment generating function (MGF)** is $M_X(t) = E(e^{tX})$ for $t \in \mathbb{R}$. If the MGF exists, then

$$M_X(0) = 1,$$
$$\frac{d^n}{dt^n} M_X(t)|_{t=0} = M_X^{(n)}(t) = E(X^n).$$

You should sanity-check that $M_X(0) = 1$ whenever you calculate an MGF.

Result 3 (MGF for sum of independent r.v.s)**.** For independent random variables, X, Y with MGFs M_X, M_Y, then $M_{X+Y}(t) = M_X(t)M_Y(t)$.

Result 4 (MGF location-scale)**.** For random variable X and scalars a, b,

$$M_{a+bX}(t) = e^{at} M_X(bt)$$

since $M_{a+bX}(t) = E(e^{t(a+bX)}) = e^{at}E(e^{btX})$.

Remark 5. A distribution is uniquely determined by any of the following:

1. PMF (common for discrete),
2. PDF,
3. CDF (common for continuous),
4. MGF, or
5. matching to a named distribution (common).

1.2 Poisson processes

Definition 6 (Poisson process)**.** Consider a problem similar to Blissville/Blotchville, where $T_1, T_2, \ldots,$ represent the arrival times of busses (the amount of time from when we started waiting to when each bus arrives). Then the bus arrival process is a **Poisson process** with rate λ if it satisfies the following conditions:

1. For any interval in time of length $t > 0$, the number of arrivals in that interval is distributed $\text{Pois}(\lambda t)$.
2. For any non-overlapping (disjoint) intervals of time, the number of bus arrivals are independent.

This applies for any "arrival process" where $T_1, T_2, \ldots$ correspond to arrival times.

☣ **7** (Poisson process units)**.** Pay attention to units: λ is a rate. So if λ has units of arrivals per hour, then t should have units of hours.

Result 8 (Inter-arrival times)**.** In a Poisson process with rate λ, the inter-arrival times (the time for the first arrival, T_1, and the times between consecutive arrives $T_2 - T_1, T_3 - T_2, \ldots$) are each independently distributed

$$T_1, T_2 - T_1, T_3 - T_2, \ldots \overset{i.i.d.}{\sim} \text{Expo}(\lambda).$$

☣ **9.** Additionally note that $T_2, T_3, \ldots,$ are *not* exponentially distributed. In fact, they follow Gamma distributions (which we will introduce soon): $T_n \sim \text{Gamma}(n, \lambda)$.

Result 10 (Count-time duality)**.** Fix a time $t > 0$. Let N_t be the number of arrivals in the time interval $[0, t]$, and let T_n be the arrival time of the n-th arrival. Then

$$(T_n > t) = (N_t < n).$$

1.3 Marginal, Conditional, and Joint Distributions

Definition 11 (Marginal, conditional, and joint distributions)**.** Consider two random variables X, Y.

	Joint	Marginal	Conditional
Distribution	(X, Y)	X	$X \mid Y = y$
PMF	$P(X = x, Y = y)$	$P(X = x)$	$P(X = x \mid Y = y)$
CDF	$P(X \leq x, Y \leq y)$	$P(X \leq x)$	$P(X \leq x \mid Y = y)$

For example, $P(X|Y = y)$ is a marginal PMF. All of these apply if we flip X and Y, and PDFs follow analogously from PMFs.

Result 12 (Marginalization)**.** If we know the joint distribution of random variables (X, Y), then we can find the marginal distribution of X (and analogously, Y) by LOTP:

$$P(X = x) = \sum_y P(X = x, Y = y), \qquad X, Y \text{ discrete.}$$
$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y), \qquad X, Y \text{ continuous.}$$

☣ **13.** Note that marginal distributions of X and Y are not sufficient (not enough information) to find the joint distribution of X, Y.

Result 14 (Joint from marginal and conditional)**.** If we know the marginal distribution of X and the conditional distributions $Y|X = x$ for any x, then we can find the joint distribution of (X, Y) by factoring out our probability:

$$P(X = x, Y = y) = P(X = x)P(Y = y|X = x), \qquad X, Y \text{ discrete.}$$
$$f_{X,Y}(x, y) = f_X(x) f_{Y|X=x}(y), \qquad X, Y \text{ continuous.}$$

Definition 15 (Independence of random variables)**.** Random variables X, Y are **independent** if for all x and y, any of the following hold (they imply each other, if valid):

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq Y) = P(X \leq x)P(Y \leq Y) = F_X(x)F_Y(y), \qquad \text{CDFs for any } X, Y.$$
$$P(X = x, Y = y) = P(X = x)P(Y = y), \qquad \text{PMFs for discrete } X, Y.$$
$$f_{X,Y}(x, y) = f_X(x) f_Y(y), \qquad \text{PDFs for continuous, } X, Y.$$

Result 16 (2D LOTUS)**.** Let X, Y be random variables with known joint distribution. For g : support$(X) \times$ support$(Y) \to \mathbb{R}$, LOTUS extends to 2 dimensions (or analogously for any larger dimensions) to give

$$E(g(X, Y)) = \begin{cases} \sum_x \sum_y g(x, y) P(X = x, Y = y), & X, Y \text{ discrete} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dx dy, & X, Y \text{ continuous.} \end{cases}$$

1.4 Covariance and correlation

Definition 17 (Covariance)**.** The **covariance** of random variables X, Y is

$$\text{Cov}(X, Y) = E\left([X - EX]\,[Y - EY]\right)$$

where EX is shorthand for $E(X)$. Equivalently,

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

Definition 18 (Correlation)**.** The **correlation** of random variables X, Y is

$$\begin{aligned} \text{Corr}(X, Y) &= \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \\ &= \frac{\text{Cov}(X, Y)}{\text{SD}(X)\text{SD}(Y)}, \end{aligned}$$

where $\text{SD}(X) = \sqrt{\text{Var}(X)}$ is the standard deviation of X. Equivalently, we first standardize X and Y, then find their covariance:

$$\text{Corr}(X, Y) = \text{Cov}\left(\frac{X - E(X)}{\text{SD}(X)}, \frac{Y - E(Y)}{\text{SD}(Y)}\right).$$

Definition 19. X and Y are

- **positively correlated** if $\text{Corr}(X, Y) > 0$,
- **negatively correlated** if $\text{Corr}(X, Y) < 0$,
- **uncorrelated** if $\text{Corr}(X, Y) = 0$.

Since correlation and covariance have the same sign, this also applies for positive/negative/zero covariance.

Result 20 (Properties of covariance: see page 327 in Blitzstein & Huang for full list)**.** Let X, Y, W, Z be random variables, as well as those of the form $X_1, X_2, \ldots,$.

- If X, Y are independent, then $\text{Cov}(X, Y) = 0$ (so X, Y are uncorrelated).
- $\text{Cov}(X, X) = \text{Var}(X)$.
- $\text{Var}(\sum_i X_i) = \sum_i \text{Var}(X_i) + \sum_{i<j} 2\text{Cov}(X_i, X_j)$.
  - This can be especially useful for finding the variance of a sum of indicators.
- $\text{Cov}(X + Y, W + Z) = \text{Cov}(X, W) + \text{Cov}(X, Z) + \text{Cov}(Y, W) + \text{Cov}(Y, Z)$.
- $\text{Cov}(aX, bY) = ab\text{Cov}(X, Y)$.

The last two properties are referred to as **bilinearity**.

Result 21 (Properties of correlation)**.** Let X, Y be random variables.

- If X, Y are independent, then $\text{Corr}(X, Y) = 0$ (so X, Y are uncorrelated)
- $-1 \leq \text{Corr}(X, Y) \leq 1$.

☣ **22** (Uncorrelated does NOT imply independent)**.** In the previous two results, we noted independent random variables have zero correlation and zero covariance. However, the converse does not apply: uncorrelated random variables are not necessarily independent.

2 Practice Problems

1. (*Example 6.6.1, Blitzstein & Hwang*) By pattern-matching to Taylor series of the exponential function, find the MGF of $X \sim \text{Pois}(\lambda)$. Recall that the support of X is $\{0, 1, 2, \ldots\}$ where, for k in the support, the PMF is

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!},$$

and the Taylor series for the exponential function is

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

2. *(inspired by Problem 4 on 2023 Stat 211 Pset # 3)* Suppose $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} \text{Expo}(\lambda)$. Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ be the **sample mean** of the random variables.

   (a) What is the joint PDF of $(X_1, \ldots, X_n)$?

(b) Show that the conditional distribution of $(X_1, \dots, X_n)$ given $\bar{X} = x$ is uniform across a region of $\mathbb{R}^n$; you can assume $x > 0$. To do so, use Bayes' rule:

$$f_{X_1,\dots,X_n|\bar{X}=x}(x_1, \dots, x_n) = \frac{P(\bar{X} = x|X_1 = x_1, \dots, X_n = x_n) f_{X_1,\dots,X_n}(x_1, \dots, x_n)}{f_{\bar{X}}(x)}.$$

Then

- Treat the denominator is constant since $f_{\bar{X}}(x)$ does not depend on $x_1, \dots, x_n$,
- Simplify $f_{\bar{X}|X_1=x_1,\dots,X_n=x_n}(x)$ using logic,
- Write $f_{X_1,\dots,X_N}(x_1, \dots, x_n)$ in terms of x.

3. Quincy competes in the intramural coin flipping league. The season consists of 5 regular-season games. At the end of the regular season, the following scenarios can occur:

   - Quincy wins less than 3 games, in which case their season is over.
   - Quincy wins at least 3 games, in which case they qualify for the playoff bracket. Given they qualify, they can...
     - play and lose in the quarterfinal: play 1 additional game and win 0
     - win the quarterfinal but lose in the semifinal: play 2 additional games and win 1.
     - win the quarterfinal and semifinal but lose in the final: play 3 additional games and win 2.
     - win the quarterfinal, semifinal, and final: play 3 additional games and win 3.

In any game that Quincy actually plays, it has a 50% probability of winning the game. Let W be the number of games Quincy wins in its intramural coin tossing season, and let R and S be the number of regular-season and playoff wins, respectively. This means $W = R + S$.

(a) Find $E(W)$. *Hint: find the distributions of R and S and use linearity.*

(b) Write the joint PMF of (R, S):